

# 基于改进 YOLOv5 的输电线路瓷瓶检测

聂嘉琦

国家电网银川供电公司，宁夏银川，750004；

**摘要：**搭载了目标检测器的无人机在进行输电线路瓷瓶检测时。受制于边缘设备内存限制，小参数数量的目标检测模型无法保证较高的检测精度，导致在实际工业应用中的效果差强人意。针对传统 YOLOv5 目标检测模型定位精度不佳的问题，本文提出改进的 YOLOv5 算法，重构了 YOLOv5 的输出头部，输出概率分布表示的坐标。并进一步引入定位分布焦点损失，提升了定位精度。实验数据证明了本文所提出改进算法切实有效地提升了检测精度。

**关键词：**YOLOv5；输电线路瓷瓶；目标检测

**DOI：**10.69979/3060-8767.25.12.024

## 1 引言

搭载了故障检测系统的无人机或智能巡检机器人实现了对输电线路的日常巡检，可以有效定位分析电气设备潜在的隐患<sup>[1-4]</sup>。航拍输电线路图中常出现背景复杂，瓷瓶较小的情况，在输电线路巡检等资源受限的场景中，瓷瓶的精准定位问题凸显了兼具轻量化与高精度检测模型的必要性。

## 2 YOLOv5 目标检测算法

YOLOv5 在检测速度与精度之间实现了良好平衡，在复杂工业场景下表现稳定并获得广泛应用。因此本文将其作为基准模型。

在模型结构上，YOLOv5 由输入端、骨干网络、颈部网络和检测头构成。输入端引入 Mosaic 进行数据增强。在骨干网络，YOLOv5 采用堆叠平行金字塔融合（SPPF）结构，该模块是空间金字塔池化（SPP<sup>[5]</sup>）的改进，通过多尺度特征融合增强模型对空间信息的提取能力。在颈部网络，YOLOv5 引入特征金字塔网络（FPN<sup>[6]</sup>）与路径聚合网络（PANet<sup>[7]</sup>）。FPN 自顶向下将深层的强语义信息与浅层特征融合，实现不同尺度特征图在语义表达上的增强。PANet 自底向上加强位置信息的传递与融合，增加网络的预测精准度。

## 3 改进的 YOLOv5 算法

在目标检测中，边界框坐标结构化表示方法主要有两种：左上右下坐标  $\{x,y,x,y\}$ ，中心点的坐标与宽高  $\{x,y,w,h\}$ 。在边界模糊等比较复杂的场景中，结构化表示具有很强的不确定性，无法精确地表达出物体的位置。

针对上述问题。本文将坐标表示改进为概率分布的形式，利用 Softmax 函数将边界框的坐标转化为离散的

概率分布回归，并引入分布焦点损失（DFL），约束模型在训练过程中拟合真实位置区域坐标。

### 3.1 概率分布表示边界框

General Focal Loss<sup>[8]</sup>中提出了使用离散概率分布来表示边界框坐标。消除边界框位置的不确定性。

设  $e \in B$  表示边界框的一条边，它的值可以表示为如下形式：

$$\hat{e} = \int_{e_{\min}}^{e_{\max}} xP(x)dx, e \in B(3-1)$$

其中  $x$  是回归坐标，其范围是  $[e_{\min}, e_{\max}]$ ， $P(x)$  是回归坐标对应的概率。将连续的回归区间  $[e_{\min}, e_{\max}]$  转换成离散变量的形式： $e = [e_1, e_2, \dots, e_n]^T \in R^n$ ，其中  $e_1 = e_{\min}$ ， $e_n = e_{\max}$ 。通过 Softmax 函数对离散的坐标区间进行回归，得到归一化的概率分布。直观地反映坐标值落在各个子区间的可能性，得出边界框的概率化表示。

### 3.2 改进的 YOLOv5 头部

针对 YOLOv5 定位边界框表示的质量问题，本文对 YOLOv5 的 Detect 层进行了重构。改进后的模型头部结构如图 3-1 所示：

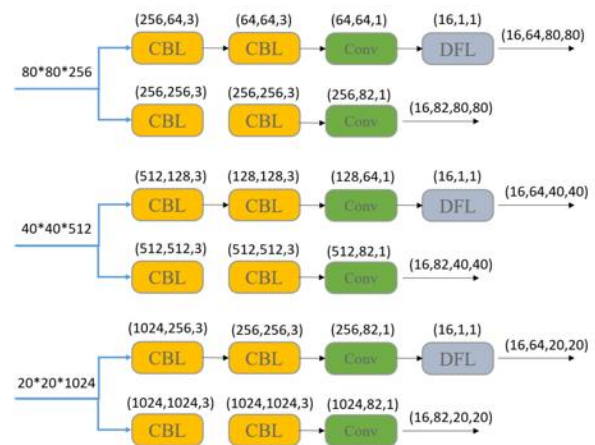


图 3-1 YOLOv5 改进 head 结构

原模型使用单一卷积层同时预测边界框位置与类别信息；改进后采用两个独立的分支完成上述任务。其中，位置预测利用 Softmax 回归坐标的概率分布。其中 CBL 模块由  $3 \times 3$  卷积、批量归一化层以及 SiLU 激活函数构成；DFL 模块由  $1 \times 1$  卷积构成，负责分布积分计算。进行概率分布表示时，取区间为  $[0, \text{reg\_max}]$ ， $\text{reg\_max}$  表示坐标离散化的最大值。预测得到的边界框标签  $\{l, r, t, b\}$  除以特征图的缩放尺度归一化后，其数值映射到  $[0, \text{reg\_max}]$  的子区间  $[i, i+1]$ 。即模型预测边界框位置最可能出现的概率区间。即概率分布  $p(i), p(i+1)$  应该更大。通常选取  $\text{reg\_max}$  值为 15，4 个方向上共预测 64 个概率分布得分。

### 3.3 改进 YOLOv5 损失函数

由于不同的概率分布可以积分得到相同的坐标表示。这导致模型在不加约束的情况下难以确定应该去拟合哪种类型的分布形状。对模型的训练中产生负面作用。

针对上述提到的问题，本文引入定位分布焦点损失 DFL Loss, DFL loss 在 Focal Loss 的基础上发展而来，使用类似于交叉熵函数的形式，使得真实标签附近位置上的概率分布数值更大，训练时模型可以更多地关注在真实标签处，有效地解决概率分布表示的不唯一性问题。DFL loss 的公式如下。

$$\text{DFL}(p(i), p(i+1)) = -[(y_{i+1} - y) * \log(p(i)) + (y - y_i) * \log(p(i+1))] \quad (3-2)$$

其中， $y$  是真实标签坐标值通过 Softmax 函数转换后得到的对应概率分布， $y_i, y_{i+1}$  是与  $y$  相邻的离散区间值。 $p(i)$  表示模型输出的预测概率处于区间  $y_i$  的概率。

当  $p(i)$  和  $p(i+1)$  的取值较大时，表示模型对于真实概率区间的预测较为准确，损失值也会减小。

## 4 实验结果与分析

### 4.1 数据集

本文采用真实航拍输电线路瓷瓶图片进行模型的训练及验证。将标注后的样本按 8:1:1 的比例分配训练集、验证集和测试集。数据集可视化图像如图 4-1 所示：



图 4-1 数据集可视化图像

### 4.2 实验结果及对比

改进前后的 YOLOv5 推理检测效果对比如图 4-2 所示。

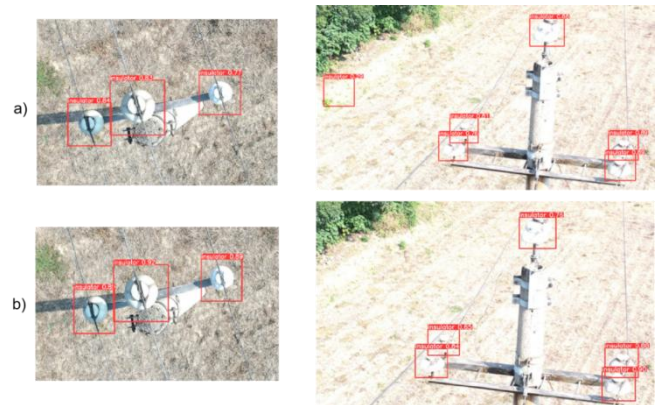


图 4-2 改进前后检测效果对比 a) 改进前；b) 改进后

从图中可以看出，改进后的 YOLOv5 排除了误检的情况。对于瓷瓶的置信度判断也更加精准，证明了本文所提改进方法的有效性。

本文在航拍输电线路瓷瓶数据集上进行训练后得到的平均检测精度如表 4-2 所示。

表 4-2 改进前后 YOLOv5 精度比较

模型	mAP_0.5:0.95
改进 YOLOv5	75.80%
原 YOLOv5	72.85%

在训练过程中的改进前后 YOLOv5 识别精度曲线如图 4-3 所示。

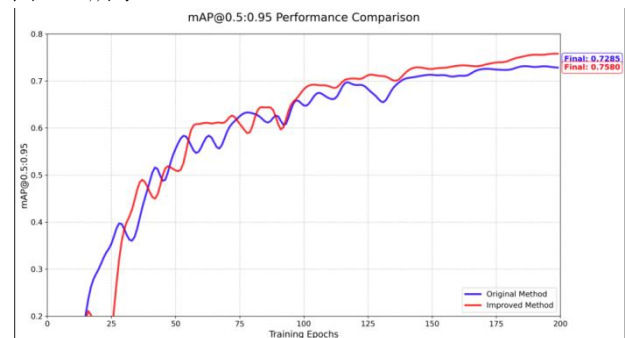


图 4-3 改进前后精度对比曲线

改进后的 YOLOv5 检测精度提高了 2.9 个百分点，在 200 个 epoch 的训练中，改进后 YOLOv5 加速收敛的现象表明其边界框预测能更快速地关注到真实标签区域，体现了模型性能的提升。

## 5 结论

本文提出了一种 YOLOv5 模型的改进方案，使用概率分布表示替代传统的边界框结构化表示，并引入了定位分布焦点损失（DFL）加强模型对真实标签区域的关

注。对比实验表明该改进模型在检测精度和收敛速度上均优于原模型，验证了改进的有效性。

### 参考文献

- [1]宋思齐. 高压电气设备红外检测技术研究[D]. 武汉: 华中科技大学, 2016.
- [2]Dun Nan Liu Rui Hou, Wen Zhuo Wu, Jing Wen Hua , Xuan Yuan Wang, Bo Pang Research on infrared image enhancement and segmentation of power equipment based on partial differential equation[J]. Journal of Visual Communication and Image Representation. 2019. 64.
- [3]X. Jianbo, K. Chen, C. Tian, F. Xiaohua and 2. Longwu, "Study on the Intelligent Analysis Method of the Infrared Detection of Electrical Equipment," 2018 3rd International Conference on Smart City and Systems Engineering (ICSCSE) Xiamen, China, 2018, pp. 500-504.
- [4]吴俊杰. 无人机智能巡视系统在变电站巡检中的

应用 [J]. 农村电气化, 2023, (10): 21-23+88. DOI:10.13882/j.cnki.ncdqh.2023.10.006.

[5]He K, Zhang X, Ren S, et al. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2015, 37(9):1904-1916.

[6]Lin T Y, Dollar P, Girshick R, et al. Feature pyramid networks for object detection[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2017: 2117-2125.

[7]Liu S, Qi L, Qin H, et al. Path aggregation network for instance segmentation[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 8759-8768.

[8]LI X, WANG W, WU L, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection[A]. 2020.